BRIEF REPORT

Temperature Screening for SARS-CoV-2 in Nursing Homes: Evidence from Two National Cohorts

Kevin W. McConeghy, PharmD, MS,[*†] *Elizabeth White, APRN, PhD,*[†] *Orestis A. Panagiotou, MD, PhD,*[†] *Christopher Santostefano, RN, BSN,*[†] *Christopher Halladay, ScM,*[*] *Richard A. Feifer, MD MPH,*[‡] *Carolyn Blackman, MD,*[‡] *James L. Rudolph, MD,*[*†§] *Vince Mor, PhD,*[*†] *and Stefan Gravenstein, MD, MPH*[*†§]

BACKGROUND/OBJECTIVES: Infection screening tools classically define fever as 38.0°C (100.4°F). Frail older adults may not mount the same febrile response to systemic infection as younger or healthier individuals. We evaluate temperature trends among nursing home (NH) residents undergoing diagnostic SARS-CoV-2 testing and describe the diagnostic accuracy of temperature measurements for predicting test-confirmed SARS-CoV-2 infection.

DESIGN: Retrospective cohort study evaluating diagnostic accuracy of pre–SARS-CoV-2 testing temperature changes.

SETTING: Two separate NH cohorts tested diagnostically (e.g., for symptoms) for SARS-CoV-2.

PARTICIPANTS

Veterans residing in Veterans Affairs (VA) managed NHs and residents in a private national chain of community NHs.

MEASUREMENTS: For both cohorts, we determined the sensitivity, specificity, and Youden's index with different temperature cutoffs for SARS-CoV-2 polymerase chain reaction results.

RESULTS: The VA cohort consisted of 1,301 residents in 134 facilities from March 1, 2020, to May 14, 2020, with 25% confirmed for SARS-CoV-2. The community cohort included 3,368 residents spread across 282 facilities from February 18, 2020, to June 9, 2020, and 42% were confirmed for SARS-CoV-2. The VA cohort was younger, less White, and mostly male. A temperature testing threshold of 37.2°C has better sensitivity for SARS-CoV-2, 76% and 34% in the VA and community NH, respectively, versus 38.0°C with 43% and 12% sensitivity, respectively.

CONCLUSION: A definition of 38.0°C for fever in NH screening tools should be lowered to improve predictive accuracy for SARS-CoV-2 infection. Stakeholders should carefully consider the impact of adopting lower testing thresholds on testing availability, cost, and burden on staff and residents. Temperatures alone have relatively low sensitivity/specificity, and we advocate any threshold be used as part of a screening tool, along with other signs and symptoms of infection. **J Am Geriatr Soc 00:1-5, 2020.**

Keywords: nursing homes; aged 80 and older; temperature; COVID-19; SARS-CoV-2

Nursing home (NH) residents with SARS-CoV-2 infection have the highest mortality rates from the COVID-19 global pandemic. NH residents account for an estimated 45% of all COVID-related deaths in the United States and make up only .6% of the total U.S. population.[1] Early identification of SARS-CoV-2 by symptomatic screening of staff and residents is critical. The Centers for Disease Control and Prevention (CDC) recommends temperature and symptom-based screening for COVID-19 on admission and at least daily for all residents.[2] Fever is a common symptom of SARS-CoV-2 infection, but definitions for febrile illness lack precision and are not widely accepted.[3-5] One definition of fever is a temperature above 100.0°F or *possibly* two above 99.0°F. Previous screening tools have

From the *Department of Veterans Affairs, Center on Innovation in Long-Term Services and Supports, Providence VA Medical Center, Providence, Rhode Island; †Department of Health Services, Policy and Practice, School of Public Health, Brown University, Providence, Rhode Island; ‡Genesis HealthCare, Kennett Square, Pennsylvania; and the §Division of Geriatrics and Palliative Care, Warren Alpert Medical School of Brown University, Providence, Rhode Island.

Address correspondence to Kevin W. McConeghy, PharmD, MS, Providence VA Medical Center, Department of Health Services, Policy and Practice, Brown University School of Public Health, Providence, RI. E-mail: kevin_mcconeghy@brown.edu; kevin.mcconeghy@va.gov. Twitter: @KevFromRI; Stefan Gravenstein, MD, MPH, Division of Geriatrics and Palliative Care, Warren Alpert Medical School of Brown University, Providence, RI. E-mail: stefan_gravenstein@brown.edu

DOI: 10.1111/jgs.16876

0002-8614/20/$15.00

used 38.0°C and higher.[6,7] However, frail older adults may be unable to mount the same temperature response to systemic infection as young healthy individuals. A so-called normal baseline temperature for NH residents is not well described nor is an age-specific definition of fever. A previous report by our research team highlighted the poor sensitivity of 38.0°C for identifying SARS-CoV-2 infection in a cohort of U.S. veterans undergoing facility-wide testing.[8] In this report we expand those findings with two separate NH populations: (1) a cohort of veterans residing in Veterans Affairs (VA) NHs, and (2) a cohort of community NH residents from a large private multistate nursing home chain. We describe the diagnostic accuracy of temperature cutoffs and hypothesized that the traditional cutoff of 38.0°C (100.4°F) has poor sensitivity for screening SARS-CoV-2 positive cases.

METHODS

This was a retrospective cohort study using clinical electronic health record (EHR) data from two separate cohorts. The first cohort included U.S. veterans receiving long-term care within the VA system at 134 NHs, also known as community-living centers, nationwide. The second cohort included residents of 369 private community NHs owned by a large multistate provider of post-acute and long-term care. The research team had access to detailed EHR data for both cohorts. The VA cohort relied on EHR data from March 1 to May 14, 2020; the community cohort included data from February 18, 2020, through June 9, 2020. We excluded residents not tested, those tested before NH admission, as part of sweep testing, or those missing temperature readings.

During this pandemic period, in both cohorts temperature was documented at least daily by nursing staff. For those with multiple temperature readings per day, we evaluated the daily maximum temperature (T-max).

To assess the diagnostic performance of different temperature cutoffs, we computed the sensitivity ("true positives"), specificity ("true negatives"), Youden's index ([sensitivity + specificity] − 1), positive predictive value, negative predictive value, area under the receiver operating characteristic curve (AUROC), and 95% confidence intervals. Youden's index for a given cutoff can be used to balance sensitivity/specificity and corresponds to the point on an ROC curve with most vertical distance from the 45-degree diagonal line (an uninformative AUROC of .5).[9] We evaluated temperatures from 36.5°C to 38.3°C and reported the diagnostic accuracy for temperature cutoffs of 37.2°C, 37.5°C, 38.0°C, and the temperature cutoff that maximized the Youden statistic.

RESULTS

A Consolidated Standards of Reporting Trials diagram in Supplementary Figure S1 shows the inclusion and exclusion allocation for our study sample. The VA NH cohort included 1,301 residents in 140 facilities from March 1, 2020, to May 14, 2020, with 25% positive for SARS-CoV-2. After excluding the untested, those tested without symptoms of disease, and missing temperature readings, the community nursing home (CNH) cohort included 3,368 residents spread across 282 facilities from February 18, 2020, to June 9, 2020, with 42% positive for SARS-CoV-2.

The VA cohort was mostly male and, compared with the community cohort, was younger, with a larger proportion of Black residents (Table 1). Both cohorts had similar baseline temperatures (36.8°C and 36.9°C, respectively) before testing. In both cohorts, residents confirmed to have SARS-CoV-2 were more likely to be Black, had higher rates of dementia, and higher recorded temperatures. In the VA cohort, more had congestive heart failure (23.6% vs 36.2%) than residents testing negative. In the CNH cohort, residents had more hypertension diagnoses, 81.4% versus 78.2%. Only about 20% of CNH patients had a 24-hour T-max of at least 38.0°C before testing, whereas 45% of VA residents did.

Figure 1 shows the percentage of residents exceeding a given maximum temperature relative to the date of testing in both cohorts. Of 3,314 residents who were tested in CNHs, 971 (29%) had a temperature of 37.2°C or higher reported in the 24 hours preceding testing. Of the 1,301 residents in the VA NHs, 535 (42%) had a temperature of 37.5°C or higher reported in the 24 hours before testing. Neither cohort's SARS-CoV-2 positive residents had an average T-max above 38.0°C in this period. Figure 2 presents the ability for discrete temperature changes to discriminate results that confirmed positive or rejected negative SARS-CoV-2 infection. Overall temperature measurements had more diagnostic value in VA NHs versus CNHs. The cutoff that maximizes sensitivity and specificity (highest Youden's statistic) was 37.3°C (99.5°F) in the VA cohort and 37.1°C (99.0°F) in the CNHs. A cutoff of 38.0°C performs poorly. Using the CDC recommendation of two temperatures above 37.2°C/99.0°F reduces the sensitivity/specificity (Youden's index) versus a single 37.2°C measurement, .71/.61 (.32) for VA residents and .21/.90 (.11) for CNH residents. ROC curves are also given in Supplementary Figure S2.

DISCUSSION

This report describes temperature changes before testing for SARS-CoV-2 in symptomatic NH residents. Overall, we find that classic definitions of fever (38.0°C [100.4°F]) lack sensitivity and are insufficient to trigger SARS-CoV-2 testing for most NH residents who have SARS-CoV-2 infection. A temperature threshold for triggering SARS-CoV-2 testing appears optimal at approximately 37.2°C in terms of balancing sensitivity and specificity but with low overall accuracy.

Youden's statistic represents the overall diagnostic accuracy of a given cutoff and is the point on the line of an ROC curve furthest from chance (i.e., AUROC of .5). Overall, this statistic is highest at cutoffs of 37.1°C and 37.3°C for community and VA homes, respectively, but overall accuracy is modest at .19 to .31. A value of 1 would be a perfectly accurate test, and 0 would be no better than random chance.

Frail older adults may not mount febrile responses to infection. Even assuming 37.0°C/98.6°F as a normal baseline temperature may not be appropriate.[10] However, clinical providers confronted with subclinical temperature

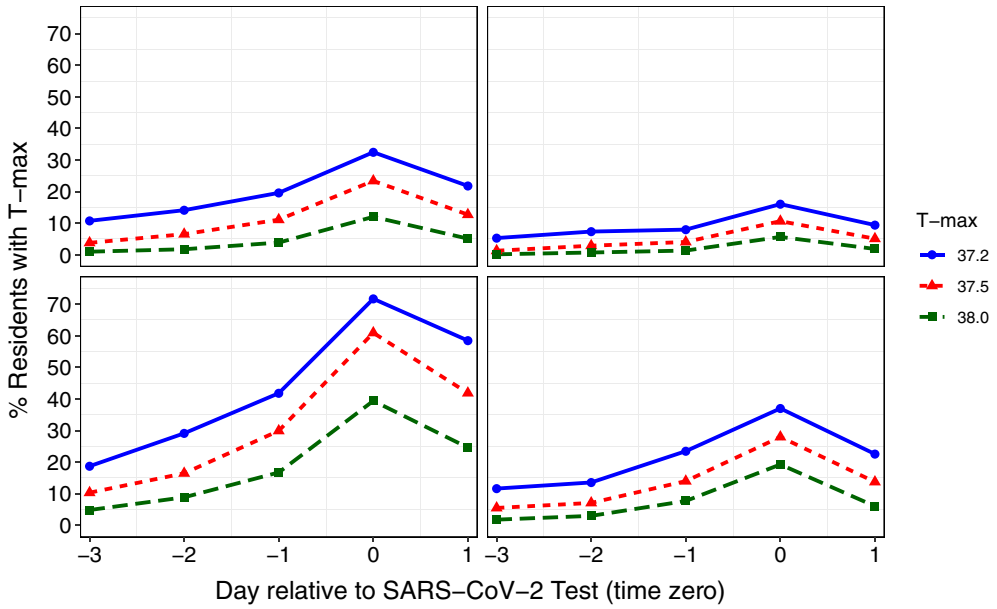


Figure 1. Percentage of residents meeting different temperature thresholds relative to SARS-CoV-2 testing. Day 0 is the date of SARS-CoV-2 test. T-max, maximum temperature in Celsius observed in a 24-hour period. Top panel shows community nursing homes; bottom panel shows Veterans Affairs nursing homes.

elevations do not have sufficient evidence to ground their clinical decision-making. Based on the accrued evidence, we recommend that NH screening tools used to aid in testing for SARS-CoV-2 lower the threshold for febrile illness definition from 38.0°C (100.4°F) to at least 37.5°C (99.5°F), if not 37.2°C (99.0°F). We favor 37.2°C, as a simple rule of thumb for those using the Fahrenheit system (rounds to 99.0°F). Without any other clinical information, this substantially improves the sensitivity for a positive SARS-CoV-2 result.

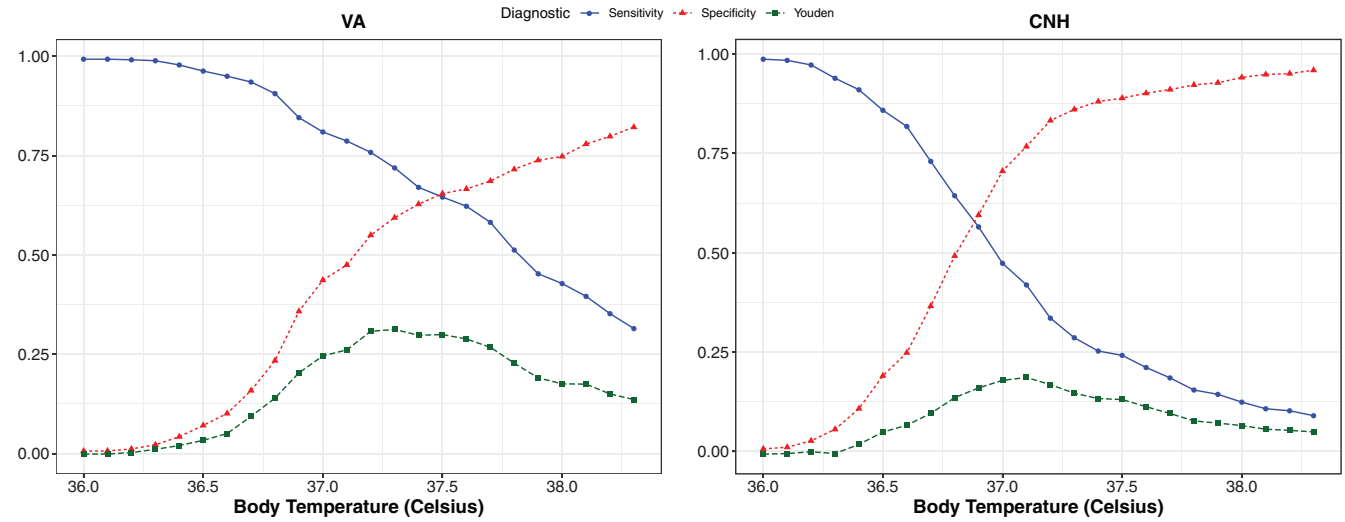
Each institution should carefully consider whether to use 37.2°C or some other cutoff. A more sensitive threshold (e.g., 37.2°C) may be favored by providers' recognition that identifying early cases could indicate outbreaks and thus may favor strategies to facilitate more testing. At the same time, more testing could limit availability of tests and be costly. Also, if residents are isolated/quarantined while waiting for results, this could further burden staff, residents, and family during a stressful time. The community prevalence of SARS-CoV-2 should also be considered. If community prevalence is low, and so pretest probability of SARS-CoV-2 is low, a lower temperature threshold will lead to more testing with little benefit. However, when community prevalence is high or there is an ongoing outbreak in

Table 1. Demographic and Clinical Characteristics of SARS-CoV-2 Tested Residents

	Veteran affairs nursing homes			Community nursing homes		
	SARS-CoV-2 (+), n = 330	SARS-CoV-2 (−), n = 971	*P* value	SARS-CoV-2 (+), n = 1,239[a]	SARS-CoV-2 (−), N = 1,657[a]	*P* value
Variable						
Age, y, mean (SD)	75.0 (10.4)	73.5 (10.8)	.02	79.1 ± 10.8	76 ± 12.6	<.01
Female, No. (%)	<10 (–)	38 (3.9)	.10	776 (62.6)	1,040 (62.8)	.92
White, No. (%)	207 (62.7)	698 (71.9)	<.01	985 (79.5)	1,343 (81.1)	.20
Black, No. (%)	98 (29.7)	209 (21.5)	<.01	189 (15.3)	194 (11.7)	<.01
Diabetes mellitus, No. (%)	133 (40.3)	433 (44.6)	.20	498 (40.2)	675 (40.7)	.70
Hypertension, No. (%)	253 (76.7)	723 (74.5)	.47	1,009 (81.4)	1,295 (78.2)	.05
CHF, No. (%)	78 (23.6)	351 (36.2)	<.01	329 (26.6)	447 (27.0)	.76
ADRD, No. (%)	236 (71.5)	544 (56.0)	<.01	658 (53.1)	795 (48.0)	<.01
Baseline temperature, mean (SD)	36.9 (.30)	36.8 (.31)	<.01	36.8 ± .19	36.8 ± .18	.48
24 h before test						
T-max, mean (SD)	37.9 (.82)	37.4 (.80)	<.01	37.3 ± .67	37.0 ± .55	<.01
T-max, ≥38.0	149 (45.4%)	225 (23.7%)	<.01	252 (20.3)	160 (9.7)	<.01
Two 37.2°C/99°F temperatures	233 (70.6%)	379 (39.0%)	<.01	185 (14.9)	107 (6.5)	<.01

Abbreviations: ADRD, Alzheimer's disease and related dementia; baseline temperature, mean maximum daily temperature up to 14 days before date of test; CHF, congestive heart failure; SARS-CoV-2, result from polymerase chain reaction test; SD, standard deviation; T-max, maximum temperature in 24 hours before test.

[a]Of the 3,368, 2,896 had complete covariate information.



Cutoff	Se	Sp	Y	PPV	NPV	AUC
37.2	.76 (.73, .78)	.55 (.53, .57)	.31 (.27, .35)	.40 (.38, .42)	.85 (.84, .87)	.62 (.61, .64)
37.5	.65 (.62, .67)	.65 (.64, .67)	.30 (.26, .34)	.42 (.40, .44)	.83 (.81, .84)	.62 (.61, .64)
38.0	.43 (.40, .46)	.75 (.73, .76)	.18 (.13, .22)	.40 (.37, .42)	.77 (.76, .78)	.58 (.57, .60)
37.3	.72 (.69, .74)	.59 (.58, .61)	.31 (.27, .35)	.41 (.39, .43)	.84 (.83, .86)	.63 (.61, .64)

Cutoff	Se	Sp	Y	PPV	NPV	AUC
37.2	.34 (.31, .36)	.83 (.81, .85)	.17 (.13, .21)	.60 (.56, .63)	.63 (.61, .65)	.58 (.57, .60)
37.5	.24 (.22, .27)	.89 (.87, .90)	.13 (.09, .17)	.62 (.57, .66)	.61 (.59, .63)	.57 (.55, .58)
38.0	.12 (.11, .14)	.94 (.93, .95)	.07 (.04, .09)	.61 (.55, .67)	.59 (.57, .61)	.53 (.52, .54)
37.1	.42 (.39, .45)	.77 (.75, .79)	.19 (.14, .23)	.57 (.54, .60)	.64 (.62, .66)	.59 (.58, .61)

Figure 2. Diagnostic accuracy of different temperature cutoffs. The diagnostic accuracy of temperature cutoffs in the 24 hours preceding the SARS-CoV-2 test in community nursing homes (CNHs) and VA NHs. Values are reported with 95% confidence intervals. AUROC, area under the receiver operating characteristic curve; NPV, negative predictive value; PPV, positive predictive value; Se, Sensitivity, Sp, specificity; Y, Youden's index. The fourth row of the table is the temperature cutoff with maximum diagnostic accuracy. [Color figure can be viewed at wileyonlinelibrary.com]

an NH, a lower threshold of 37.2°C could be a useful part of screening, increasing the ability to diagnose cases early.

Our findings are supported by the analysis of two separate NH populations. Both cohorts were constructed and analyzed separately, using different definitions of diagnostic testing, data collection methods, and analytical programmers. In the case of the CNHs, the method of temperature measurement was highly variable by facility, and prior research has demonstrated that different body sites yield inconsistent measures of core body temperature. Each cohort came to slightly different conclusions on the best cutoff to use (37.1°C vs 37.3°C). Also, in the VA cohort, temperature was overall a more sensitive/specific indicator for infection. This may be due to differences in age or sex; the VA cohort was younger and male. The VA NHs consistently use oral thermometers versus CNHs that use varying methods. A second limitation (or strength) was that the definition of "diagnostic" testing varied by cohort. In CNH this included tested within 2 days of a new-onset symptom (e.g., cough). The VA NHs did not have a regular screening tool used in this period, and so tests performed sporadically (not part of a facility-wide sweep) were assumed to be for "symptom."

The challenge facing NHs in the United States and internationally is unprecedented. Researchers and clinician partners can help limit the burden of the COVID-19 pandemic through effective infection control and rapid identification of cases. Our work supports this effort by challenging the use of classical definitions of fever and providing real-world evidence of the clinical utility of a lower threshold of 37.2°C (99.0°F) in an NH population. We urge infectious diseases clinicians and front-line providers to consider the impact of aging and frailty on immune response and highlight the need for population-specific recommendations when evaluating febrile illness.

In conclusion, this study consisted of two national cohorts of NH residents tested for SARS-CoV-2. We described the relative inability of this population to mount a fever of 38.0°C and demonstrated a lower threshold of 37.2°C improves sensitivity, but temperatures are poor diagnostic tools on their own. We recommend NHs consider that screening tools that incorporate temperature lower their test-triggering thresholds from 38.0°C.

ACKNOWLEDGMENTS

We acknowledge and appreciate Rich Castor and Cliff Boyd at Genesis HealthCare, Jeff Hiris at Brown University, and Frank DeVone at the VA for their efforts in extracting the electronic health data used for this study.

Financial Disclosure: This work was supported by the National Institute on Aging (Grant No. 3P01AG027296-11S1; principal investigator [PI] Vincent Mor) and VA Health Services Research and Development (Grant No. C19-20-213; PI James Rudolph) from the U.S. Department of Veterans Affairs Health Services R&D Service.

Conflict of Interest: Kevin W. McConeghy reports grants from Sanofi Pasteur and Pfizer outside of the submitted work. Elizabeth White reports support from the

National Institute on Aging and employment by the PACE organization of Rhode Island. Vince Mor reports grants from Sanofi Pasteur, Pfizer, and Seqirus related to vaccination in nursing homes. He also receives fees for chairing the Scientific Advisory Committee of NaviHealth, a post-acute care company. Stefan Gravenstein reports grants and personal fees from Sanofi Pasteur and Pfizer, and consulting or speaker fees from Catapult Consultants, GlaxoSmithKline, Healthcentric Advisors, Janssen, Merck, Novartis, Pfizer, and Longeveron related to vaccines or nursing home care quality. The remaining authors have declared no conflicts of interest for this article.

Author Contributions: Kevin W. McConeghy conceived the main analytical approach, performed the primary analysis for the community nursing homes, and authored the article. Christopher Santostefano and Christopher Halladay assisted with programming and analysis. Vince Mor, Elizabeth White, Richard Feiffer, Carolyn Blackman, Orestis A. Panagiotou, and Stefan Gravenstein assisted with study design, interpretation of results, and co-authored the article. Christopher Halladay and James L. Rudolph provided the primary results for the Veterans Affairs nursing homes and co-authored the article.

Sponsor's Role: The sponsor had no role in the design, conduct, analysis, or reporting of study results. The corresponding author had final responsibility for the decision to submit for publication. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the U.S. government.

REFERENCES

1. Kaiser Family Foundation. State data and policy actions to address coronavirus. https://www.kff.org/health-costs/issue-brief/state-data-and-policy-actions-to-address-coronavirus/. Accessed June 25, 2020.
2. Centers for Disease Control and Prevention. Testing guidelines for nursing homes. https://www.cdc.gov/coronavirus/2019-ncov/hcp/nursing-homes-testing.html. Accessed June 26, 2020.
3. Zhang G, Zhang J, Wang B, Zhu X, Wang Q, Qiu S. Analysis of clinical characteristics and laboratory findings of 95 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a retrospective analysis. Respir Res. 2020;21(1):74.
4. Gou FX, Zhang XS, Yao JX, et al. Epidemiological characteristics of COVID-19 in Gansu province [in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi. 2020;41:E032.
5. Kato H, Shimizu H, Shibue Y, et al. Clinical course of 2019 novel coronavirus disease (COVID-19) in individuals present during the outbreak on the diamond princess cruise ship. J Infect Chemother. 2020;26(8):865-869.
6. Alsofayan YM, Althunayyan SM, Khan AA, Hakawi AM, Assiri AM. Clinical characteristics of COVID-19 in Saudi Arabia: a national retrospective study. J Infect Public Health. 2020;13:920-925. https://doi.org/10.1016/j.jiph.2020.05.026.
7. Frankel T. A fever is 100.4 in Ohio; it's 99.5 in Delaware: states, companies write their own rules for temperature screening in a pandemic. *Washington Post*. May 15, 2020. https://www.washingtonpost.com/business/2020/05/15/fever-screening-coronavirus/. Accessed July 6, 2020.
8. Rudolph JL, Halladay CW, Barber M, et al. Temperature in nursing home residents systematically tested for SARS-CoV-2. J Am Med Dir Assoc. 2020;21:895-899.e1. https://doi.org/10.1016/j.jamda.2020.06.009.
9. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. Am J Epidemiol. 2006;163(7):670-675.
10. Protsiv M, Ley C, Lankester J, Hastie T, Parsonnet J. Decreasing human body temperature in the United States since the industrial revolution. Elife. 2020;9:e49555.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Supplementary Figure S1: Resident inclusion/exclusion, cohort workflow.

Supplementary Figure S2: Receiver operating characteristic curves.